

I enclose a file containing the nucleotide sequence of the PRE7(IAP) as is the provisional name we have selected for this element, which is involved in posttranscriptional regulation.

We have identified elements that have homology to PRE7(IAP) from the databases. The alignment of these elements is provided. None of these elements has been characterized functionally. Our work is the first that identified these sequences as potential posttranscriptional control elements. At present, PRE7(IAP) is the only one for which we have functional data.

In the aligned sequences, our functional PRE7(IAP) is called FNC3B. Only the "core" 231 nt element containing the functional element is shown.

FNC3B

GTGGGGTGCAGGGCTAAGCACTGCACAGAGGATAGCTTGCTGTTGGCATCCTGTGGAAGG  
CACGTCTGATTGCATGAAGGTTCAAGTCAGTGCCTAGTTCCCTCCCCAGGAAAAACGACACG  
GGAGCTGGCCAAGACCTCTCTGGGTGATGAGCCTAAGGGATGGTTGTAGGGCCCT  
ATGCTTGCACACTGGGATCAGACCTCTACCTCACCCATGAGGCTTGCTT

PRE7 (IAP)

CTTCGCCATGGTAGCATAGGCTTTGCTGCAGTGGAGGCGGGACAATCTCCTCAGATT  
GGTTTGCCTCTAAAGAAATTATGCTGCTTATGCCGTGGGTGCGAGGTAAGCACT  
GCACAGAGGATAGCTTGCTGTGGCATCCTGTGGAAGGCACGTCTGATTGCATGAAGGTT  
CAGTGCCTAGTTCCCTCCCCAGGAAAACGACACGGGAGCTGCCAAGACCTCTCG  
GGTGATGAGCCTAAGGGATGGTTTGTGTAGGGCCCTATGCTTGCACACTGGGATCAG  
ACCTCTACCTCACCCATGAGGCTTGCTGCAGCAATTAAGATCTGCCATAGGTTAATT  
AACATCCTGGCTTTGATGCACCTGCCACAAG

the underlined sequences correspond to the aligned fragment  
FNC3B

The alignment of the sequences found in the database follows:

FNC3B.seq Begin:99 End:329  
 GB:X57268 Begin:1599 End:1841  
 GB:M10134 Begin:7308 End:7538  
 GB:X01172 Begin:2423 End:2668  
 GB:M12515 Begin:188 End:418  
 AE000664.seq Begin:118019 End:118264  
 GB:M18252 Begin:3418 End:3662  
 GB:M18251 Begin:1630 End:1874  
 GB:S74315 Begin:3824 End:4068  
 GB:M10062 Begin:2711 End:2955  
 GB:E00593 Begin:2711 End:2955  
 E01116.gb\_pat Begin:2711 End:2955  
 GB:E00594 Begin:3537 End:3781  
 E01117.gb\_pat Begin:4388 End:4632  
 GB:X54077 Begin:2200 End:2444  
 GB:X04120 Begin:4477 End:4721  
 GB:X97915 Begin:839 End:1083  
 GB:M17551 Begin:6474 End:6721  
 GB:U58494 Begin:6192 End:6439  
 U70139.gcg Begin:4321 End:4565

!!NA MULTIPLE ALIGNMENT 1.0  
 PileUp of: @CTE-setClg.list

Symbol comparison table: GenRunData:pileupdna.cmp CompCheck: 6876

GapWeight: 2  
 GapLengthWeight: 1

CTE-setClg.msf MSF: 251 Type: N 10:03 Check: 7329 ..

Name: FNC3B	Len:	251	Check:	6699	Weight:	1.00
Name: X57268	Len:	251	Check:	7888	Weight:	1.00
Name: M10134	Len:	251	Check:	8724	Weight:	1.00
Name: X01172	Len:	251	Check:	3612	Weight:	1.00
Name: M12515	Len:	251	Check:	112	Weight:	1.00
Name: AE000664	Len:	251	Check:	4675	Weight:	1.00
Name: M18252	Len:	251	Check:	1995	Weight:	1.00
Name: M18251	Len:	251	Check:	2423	Weight:	1.00
Name: S74315	Len:	251	Check:	1833	Weight:	1.00
Name: M10062	Len:	251	Check:	1751	Weight:	1.00
Name: E00593	Len:	251	Check:	1751	Weight:	1.00
Name: E01116	Len:	251	Check:	1751	Weight:	1.00
Name: E00594	Len:	251	Check:	1811	Weight:	1.00
Name: E01117	Len:	251	Check:	1811	Weight:	1.00
Name: X54077	Len:	251	Check:	2537	Weight:	1.00
Name: X04120	Len:	251	Check:	2537	Weight:	1.00
Name: X97915	Len:	251	Check:	3213	Weight:	1.00
Name: M17551	Len:	251	Check:	5139	Weight:	1.00
Name: U58494	Len:	251	Check:	5139	Weight:	1.00
Name: U70139	Len:	251	Check:	1928	Weight:	1.00

//

1						
FNC3B	GTGGGGTGC	AGGCTAAGC.	ACTGCACAGA	GGATAGCTT.	...GCTGT.T	GG.CATCCTG T.GGAAGGCA CGTC
X57268	GAGAGTTGTA	AGACTAAGT.	ACTGCACAGA	GATTAGTCTA	GAAGCTGT.T	GGACAGTCTC T.GAGAGGCA TGTC
M10134	AGGAGTTGCA	AGGCTAAGC.	ACTGCACAGG	AGAGG.TCTG	CGG..TATAA	CGACTTTCTC CTGGGAGATA AGTC
X01172	GAGAGTTGCA	CGGCTAAGC.	ACTGCAGTAG	AAGGGCTCTG	CGGCACATAT	GAGCCTATTTC TAGGGAGACA TGTC
M12515	GAGAGTCG.A	AGGCTAAGCA	ACTGCAATGG	AAAGGCTCTG	CGGCATATAT	GAGCCTATTTC TAGGGAGACA TGTC
AE000664	GAGAGTTGCA	CGGCTAAGC.	ACTGCAGTAG	AAAGGCTCTG	CGGCATATAT	GAGCCTATTTC TAGGGAGACA TGTC
M18252	GAGAGTTGCA	CGGCTAAGC.	ACTGCAATGG	AAAGGCTCTG	CGGCATATAT	GAGCCTATTTC TAGGGAGACA TGTC
M18251	GAGAGTTGCA	AGGCTAAGC.	ACTGCAATGG	AAAGGCTCTG	CGGCATATAT	GAGCCTATTTC TAGGGAGACA TGTC
S74315	GAGAGTTGCA	CGGCTAAGC.	ACTGCAATGG	AAAGGCTCTG	CGGCATATAT	GAGCCTATTTC TAGGGAGACA TGTC
M10062	GAGAGTTGCA	AGGCTAAGC.	ACTGCAATGG	AAAGGCTCTG	CGGCATATAT	GAGCCTATTTC TAGGGAAACA TGTC
E00593	GAGAGTTGCA	AGGCTAAGC.	ACTGCAATGG	AAAGGCTCTG	CGGCATATAT	GAGCCTATTTC TAGGGAAACA TGTC

E01116 GAGAGTTGCA AGGCTAAC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GAGCCTATT TAGGGAAACA TGTC  
 E00594 GAGAGTTGCA AGGCTAAC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GAGCCTATT TAGGGAGACA TGTC  
 E01117 GAGAGTTGCA AGGCTAAC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GAGCCTATT TAGGGAGACA TGTC  
 X54077 GAGAGTTGCA AGGCTAAC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GAGCCTATT TAGGGAGACA TGTC  
 X04120 GAGAGTTGCA AGGCTAAC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GAGCCTATT TAGGGAGACA TGTC  
 X97915 GAGAGTTGCA AGGCTAAC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GTGCCTATT TAGGGGAGACA TGTC  
 M17551 GAGAGTTGCA CGGCTAAC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GAGCCTATT TAGGGAGACA TGTC  
 U58494 GAGAGTTGCA CGGCTAAC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GAGCCTATT TAGGGAGACA TGTC  
 U70139 GAGAGTTGCA AGGCTAAC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GAGCCTATT TAGGGNGACA TGTC

101

FNC3B AGTCCCTTC .CCCCCAGGAA AAACGACACG GGAGCTGGCC AAGACC.TCT CTGGGTGA.. ....TGAGC CT.A  
 X57268 AGGGACCTTT .CCCCCAGAAA AAAGGGCACA GGAGCAGGTC AGGGTT.ACT CTGGGTAAAG ATCTGTGGC CT.G  
 M10134 GATCCCTT. .CCCCCAGAA AAAAGACATC GGA.CTGGTC AGGACTTCCT CTGGGGATAA G.....ACC CTGG.  
 X01172 AGTCCCTTC TCCCCAGGAA AAACGGCACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAA GCCTGTGAGC CT.A  
 M12515 AGTCCCTTC .CCCCCAGGCA AAACGACACG GGA.CAGGTC AGGGTT.GCT CTGGGTAAA .CCTGTAAGC CT..  
 AE000664 AGTCCCTTC TCTCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAA GCCTGTGAGC CT.A  
 M18252 AGTCCCTTC .CCCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAA GCCTGTGAGC CT.A  
 M18251 AGTCCCTTC .CCCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAA GCCTGTGAGC CT.A  
 S74315 AGTCCCTTC .CCCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAA GCCTGTGAGC CT.A  
 M10062 AGTCCCTTC .CCCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAA GCCTGTGAGC CT.A  
 E00593 AGTCCCTTC .CCCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAA GCCTGTGAGC CT.A  
 E01116 AGTCCCTTC .CCCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAA GCCTGTGAGC CT.A  
 E00594 AGTCCCTTC .CCCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAA GCCTGTGAGC CT.A  
 E01117 AGTCCCTTC .CCCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAA GCCTGTGAGC CT.P  
 X54077 AGTCCCTTC .CCCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAA GCCTGTGAGC CT.P  
 X04120 AGTCCCTTC .CCCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAA GCCTGTGAGC CT.P  
 X97915 AGTCCCTTC .CCCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAA GCCTGTGAGC CT.G  
 M17551 AGTGCCT. TCCCCAGGCA AAACGACAC. GGAGCAGGTC AGGGTTGGCT CTGGGTAAA GCCTGTGAGC CTCP  
 U58494 AGTGCCT. TCCCCAGGCA AAACGACAC. GGAGCAGGTC AGGGTTGGCT CTGGGTAAA GCCTGTGAGC CTCP  
 U70139 AGTCCCTTC .CCCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAA GCCTGTGAGC CT.P

201

251

FNC3B TG.CTTGCAC ACTGGGGATC AGACCTCTAC CTTCACCAT GAGGCTTGCT T  
 X57268 AC.ATGACAC ACTGGGGATC AGACCTCTAC CTCTACCCAC GGAGCTTGCT T  
 M10134 T...TTGCAC A.TGGGGATT TGACCTCTAT CTCCACTC.C AAAGTTGTGG G  
 X01172 T.ACCTGCAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G  
 M12515 TTACCTACAC ACTGGG.... TGACCTCTAT CT.CACTCTC ATCAATATGG T  
 AE000664 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAGTTGGG T  
 M18252 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G  
 M18251 T.ACCTACAC ACTGGGGATT TGACCTCTAT CCTCACTCTC ATTAATATGG G  
 S74315 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCAGTCTC ATTAATATGG G  
 M10062 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G  
 E00593 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G  
 E01116 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G  
 E00594 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G  
 E01117 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G  
 X54077 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G  
 X04120 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G  
 X97915 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G  
 M17551 TGACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G  
 U58494 TGACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G  
 U70139 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G

Plurality: 2.00 Threshold: 1 AveWeight 1.00 AveMatch 1.00 AvMisMatch 0.00

PRETTY of: CTE-setClg.msf(\*) 10:04 ..